DEVELOPMENT OF BRIGHT-FIELD MICROSCOPE IMAGE RESOURCE, AND IDENTIFICATION OF ROBUST FOCUS MEASURE FUNCTIONS FOR AUTOMATED CAPTURING OF IMAGES

Synopsis submitted in fulfillment of the requirement for the Degree of

DOCTOR OF PHILOSOPHY

IN

BIOINFORMATICS

BY

MOHAMMAD IMRAN SHAH
Enrollment No. 136504

Under the Supervision of

Dr. Chittaranjan Rout

Department of Biotechnology and Bioinformatics
Jaypee University of Information Technology, Waknaghat, Solan–173234
Himachal Pradesh, INDIA

July, 2017

Table of Contents Page No.

SUMMARY i

CHAPTER 1
INTRODUCTION 1-5
1.1 Tuberculosis and its Screening 1
1.2 Automated Screening 2
1.3 Problem Identification 4
1.4 Objectives of Thesis 4
1.5 Organization of the Thesis 5

CHAPTER 2
ZIEHL-NEELSEN SPUTUM SMEAR MICROSCOPY IMAGE DATABASE… 6-10
2.1 Background 6
2.2 Material and Method 7
2.3 Results and Discussion 9
2.4 Conclusion 10

CHAPTER 3
IDENTIFICATION OF ROBUST FOCUS MEASURE FUNCTIONS… 11-15
3.1 Background 11
3.2 Material and Method 12
3.3 Results and Discussion 14
3.4 Conclusion 15

CHAPTER 4
ESTABLISHMENT OF HYBRIDIZED FOCUS MEASURE FUNCTIONS… 16-20
4.1 Material and Method 16
4.2 Results and Discussion 18
4.3 Conclusion 20

CHAPTER 5
AUTOMATIC DETECTION AND CLASSIFICATION OF TUBERCULOSIS… 21-24
5.1 Material and Method 21
5.2 Results and Discussion 23
5.3 Conclusion 24

CHAPTER 6
CONCLUSION AND FUTURE WORK 25
6.1 Conclusion 25
6.2 Future Work 25

REFERENCES 26-30

PUBLICATIONS 31

SUMMARY

An efficient healthcare system always takes advantage of observation and interpretation of medical data that include narrative, textual data, numerical measurement, recorded signals, radiographs and pathological images. Medical imaging is one of the most important tools for disease diagnosis. Medical image digitization and analysis have been attracting attention in recent years due to the major societal impact of medical images in the areas of diagnosis, observation and the training of doctors. From simple chest X-rays (CXR) to pathological microscopic images (sputum smear image) require the highest levels of quality among all others. Infectious disease like tuberculosis which is caused by *Mycobacterium tuberculosis* requires radiological (CXR) and pathological (Sputum smear microscopy) tests for the effective diagnosis. However, the CXR is not a confirmatory test due to its non-specific and redundant patterns. Therefore, Ziehl-Neelsen stained conventional bright field microscopic (CM) test is the most widely used confirmatory method in low and middle income countries. However, the manual screening of tuberculosis bacilli using sputum smear CM microscopy may misdiagnose 33 to 50% of active cases due to patient load at the hospital. The majority of current issues on tuberculosis diagnosis can be addressed by incorporating automated methods. Autofocusing, auto-stitching, and image segmentation and classification are the three sequential steps in automated microscopy system for the tuberculosis screening. However, lack of unified datasets impedes the development of robust algorithms on these three domains. Keeping in view of these limitations, the proposed thesis work is based on four objectives which are described in the four different chapters (Chapter 2, 3, 4 & 5). In the 1st objective, Ziehl-Neelsen Sputum smear Microscopy image Database (ZNSM-iDB) has been developed to facilitate the development of algorithms and methods related to automated microscopy system, and it is freely available at http://14.139.240.55/znsm. In the 2nd objective, the robust focus measure functions were identified for the automated capturing of Ziehl-Neelsen (ZN) sputum smear images. In the 3rd objective, a robust hybrid focus measure function system was established as a universal method for auto-focusing. In the final objective, automatic bacilli segmentation was performed using watershed segmentation method to identify whether a patient is tuberculosis positive or negative. Accomplishment of these objectives has advanced the knowledge of automated microscopy development for tuberculosis diagnosis.

CHAPTER # 1

INTRODUCTION

The treatment of any illness is always based on the observation and interpretation of medical data [1]. Collection and interpretation of medical data into meaningful information is always primary to healthcare process whether it is early Greek literature for disease description, guideline and management, or the contemporary way of using laboratory and X-ray studies to diagnose and treat disease [1]. This medical data ranges from narrative, textual data and numerical measurement to record signals, drawing, and radiographs and pathological images [1]. Medical imaging techniques such as radiography (*e.g.* 2D radiograph, computed tomography (CT), etc.) and pathological imaging (sputum smear microscopy) are very important in disease diagnosis. Some of these medical imaging tests (like MRI, CT, etc.) are either not available or not affordable for common people in low- and middle-income countries [2], while inexpensive and easy test like Ziehl-Neelsen (ZN) stained bright filed conventional sputum smear test has higher misdiagnosis rate.

Technological advances like digitization of these imaging techniques and computer aided diagnosis (CAD) have the potential to overcome some of the challenges faced by healthcare providers in developing countries [2]. Digitization and computerized analysis of these medical imaging modalities have been attracting attention in recent years due to their major societal impact of medical images in the domains of diagnosis, observation, and the training of doctors. CAD methods can improve the accuracy and reduce the time complexity of poor test like conventional microscopy (CM) or non-confirmatory chest X-Ray. Lack of sufficient infrastructures and high burden of patients in low- and middle-income countries such as India and South Africa motivated us to contribute better disease diagnosis using image analysis techniques. To accomplish this task, we have chosen tuberculosis disease which is the leading cause of death by an infectious agent worldwide (WHO, 2015).

1.1 Tuberculosis and its Screening

Tuberculosis is ranked first alongside the human immunodeficiency virus (HIV) infection as a leading cause of death worldwide (WHO, 2015). In lower and middle-income tuberculosis-endemic countries, sputum smear microscopy test is the most widely used technique for tuberculosis diagnosis due to its accessibility, low operational cost and requirement of minimal

bio-safety standards [3]. However, the sensitivity of this test is poor and varies from 50 to 60% [3]. This test is also laborious in nature. Although several other methods such as culture test, GeneXpert, etc., are available for its diagnosis, these methods are unfeasible in developing and tuberculosis-endemic countries like India due to the longer time of diagnosis and the high cost of the equipment [4]. Sputum microscopy test is mainly performed by two different methods, namely, ZN stain conventional bright field microscopy (CM) and Auramine-O stain fluorescence microscopy (FM) [4, 5, 6]. The former is cost-effective approach and requires very low maintenance and expertise. Therefore, the most tuberculosis-affected countries in Asia and Africa use the CM test as a primary technique for tuberculosis detection [4]. The manual screening of tuberculosis bacilli may misdiagnose 33 to 50% of active cases [4] due to patient load at the hospital, especially in the cases of extra-pulmonary, pediatric or patients co-infected with HIV because of the very low bacterial load in sputum [7, 8, 9]. Therefore, the increase in sensitivity of popular, cost-effective and non-invasive smear microscopy (CM) test can bring better health outcomes against tuberculosis.

1.2 Automated Screening

The majority of current issues on tuberculosis diagnosis can be addressed by incorporating automated methods (computer aided diagnosis) that increase the sensitivity and specificity of CM images by analyzing a large number of fields of view for *Mycobacterium tuberculosis* identification [10]. Automation prevents physical and mental fatigue, and reduces the time of field of view analysis [7]. Furthermore, it allows storing and sharing of the images for review in future. Automation will help in the development of automated microscopy for ZN sputum smear images.

Autofocusing, autostitching and automated grading methods are necessary for automated microscope development; therefore, a unified and standard resource is required to compare existing methods as well as to develop new algorithms. In bright field conventional microscopy, Kinyoun-stained sputum smear bright-field microscopic images are available for autofocusing and bacilli segmentation [11], but bacilli in these images are less visible leading to poor sensitivity than those in ZN-stained images [12]. Therefore, a diverse imaging database for ZN sputum smear images is provided which assist to develop or validate automated algorithms. Autofocusing, autostitching and automated grading methods are described below:

1.2.1 Autofocusing

In a stack of images captured from a single view-field with different focuses, an image with the best average focus over entire view-field is defined as the focused one. The maximum value of the focus measure function (FMF) corresponds to the best-focused image [13]. This method would facilitate automated capturing of the best-focused image. Several autofocus algorithms have been reported in the literatures. Thirteen FMFs were employed on the stacks of FM smear images. The mid-frequency discrete cosine transform (96.67%), Vollath's autocorrelation (89%) and Tenengrad (89%) methods were found to be efficient [14]. Costa et al., [15] had evaluated nine FMFs on Kinyoun staining sputum smear slides which have lower sensitivity rate than CM images [12]. Studies have also indicated that some FMFs performed significantly better in fluorescence microscopy images [16-18] but produce an average outcome in CM. Although many FMFs were proposed for auto-focusing, their accuracies were mainly dependent on the content of images and instrument [19].

1.2.2 Autostitching

This method stitches view-fields of a smear-slide to form a mosaic or slide map. The WHO recommended analysis of 300 view-fields can be achieved faster and efficiently through automated stitching of overlapping view-fields followed by detection of bacilli in a mosaic (stitched-image) using segmentation [20].

1.2.3 Automatic Bacilli Segmentation and Grading

It is a process of segmentation and counting of bacilli either from view-field or mosaic. Pattern recognition and machine learning techniques have been used to detect the bacilli in images [4, 5, 7, 12, 21, 22], but the efficacy and scopes are limited due to their implementation on non-unified and limited datasets. Several methods have previously been proposed for bacilli detection in fluorescence [23-26] and conventional [27-29] microscopic images. Detection of bacilli from CM is more challenging than fluorescence microscopy due to artifacts, over-staining and lack of clear separation of bacilli from background. Costa et al., have developed a neural network based approach [27] which initially determine the density of background using the Hue component in the image. Set of color features were used for bacilli segmentation. Color ratio was used as an additional parameter to overcome the misclassification due to high density background. Zhai et

al. [28] converted RGB images into HSV and Lab color space to perform coarse and fine segmentation of the CM images. In 2011, Osman et al. [29] used a single-layer feed-forward neural network technique for tuberculosis bacilli detection. Most of the previously reported methods have shown difficulty in the segmentation and classification of overlapped/occluded bacilli.

1.3 Problem Identification

In this thesis, we have proposed the improvement of steps related to automated microscopy for tuberculosis diagnosis using ZN sputum smear microscopy images. Following are the limitations of existing available algorithms.

- A unified referral resource is required in public domain to check the robustness of various automated microscopy algorithms on ZN sputum smear images. Lack of unified database hindered the validation of these methods.
- Generally, performances of autofocus algorithms are dependent on different imaging conditions and modalities. Ample of autofocus algorithms have been reported in the literature for different imaging modalities, but most of them are not evaluated systemically on ZN sputum smear images.
- Several *Mycobacterium tuberculosis* bacilli segmentation methods have been reported in literature, but the most of them are unable to detect overlapping or occluded bacilli.

1.4 Objectives of Thesis

To overcome the limitations of automated microscopy domain, we have defined the following four objectives.

1. *Development of a **Z**iehl–**N**eelsen **S**putum smear **M**icroscopy image **D**ata**B**ase (ZNSM-iDB) which facilitate the validation of automated microscopy algorithms.*
2. *Identification of robust focus measure functions for the automated capturing of focused images from Ziehl-Neelsen stained sputum smear microscopy slide.*
3. *Establishment of hybridized focus measure functions as universal method for auto-focusing.*
4. *Automatic detection and classification of tuberculosis bacilli from ZN-stained sputum smear images using watershed segmentation.*

1.5 Organization of the Thesis

The thesis has been organized into distinct six chapters as follows.

In this chapter (Chapter 1), introductory part and related work of the thesis has presented. In Chapter 2, a newly developed **Z**iehl–**N**eelsen **S**putum smear **M**icroscopy image **D**ata**B**ase (ZNSM-iDB) has been presented which can assist the development and validation of automated microscopy algorithms. In Chapter 3, a systematic analysis on identification of robust focus measure functions for the automated capturing of focused images from Ziehl-Neelsen stained sputum smear microscopy slide has been discussed. In Chapter 4, a detailed study is provided on establishment of hybridized focus measure functions as universal method for auto-focusing. In Chapter 5, a study on automatic detection and classification of tuberculosis bacilli from ZN-stained sputum smear images using watershed segmentation has been presented. In final chapter (Chapter 6), conclusion of the thesis and future prospect has been presented.

CHAPTER # 2

ZIEHL-NEELSEN SPUTUM SMEAR MICROSCOPY IMAGE DATABASE (ZNSM-iDB): A RESOURCE TO FACILITATE AUTOMATED BACILLI DETECTION FOR TUBERCULOSIS DIAGNOSIS

Ziehl-Neelsen stained microscopy is the crucial bacteriological test for tuberculosis detection, but its sensitivity is poor. Tuberculosis diagnosis through automated capturing of focused image (autofocusing), stitching of view-fields to form mosaic (autostitching), and automatic bacilli segmentation (grading) is required to significantly improve the sensitivity. However, lack of unified datasets impedes the development of robust algorithms on these three domains. Therefore, Ziehl-Neelsen Sputum smear Microscopy image Database (ZNSM-iDB) has been developed, and it is freely available at http://14.139.240.55/znsm. This database contains seven categories of diverse datasets acquired from three different bright-field microscopes. Datasets related to Autofocusing, autostitching and manually segmented bacilli can be used for developing algorithms, while other four datasets are provided to streamline the sensitivity and specificity.

2.1 Background

The automated microscopy for bacilli detection requires efficient algorithms in following three domains:

i. ***Autofocusing:*** This method would facilitate automated capturing of the best-focused image. In recent years, various autofocusing algorithms have been proposed and implemented in microscopy images for diverse biological applications [13-16, 18, 19].

ii. ***Autostitching:*** This method stitches view-fields of a smear-slide to form a mosaic or slide map. The WHO recommended analysis of 300 view-fields can be achieved faster and efficiently through automated stitching. Autostitching also facilitates automatic detection of bacilli on the edge by joining half bacillus structures on the boundaries of two different view-fields.

iii. ***Automatic bacilli segmentation and grading:*** It is a process of segmentation and counting of bacilli either from view-field or mosaic. Pattern recognition and machine learning techniques have been used to detect the bacilli in images [4, 5, 7, 12, 21, 22], but

their efficacy and scopes are limited due to the implementation on non-unified and limited datasets.

Databases and tools have already provided the better diagnosis of cancer and other diseases [30, 31]. Databases were used to develop algorithms/methods, which were implemented in computer-aided diagnosis (CAD) systems to obtain the second opinion about a disease [32-34]. The databases and CADs are also effectively used for the early detection of diseases. Keeping in view of these accomplishments of databases, the Ziehl-Neelsen Sputum smears Microscopy image Database (ZNSM-iDB) has been developed (http://14.139.240.55/znsm). This database contains diverse categories of image datasets with both medium and high noise backgrounds. It possesses datasets for all three processes required for automated microscope development. Datasets were validated by various automated microscopy algorithms to establish their robustness. This database can be used to develop and evaluate efficient and robust algorithms related to automated microscopy.

2.2 Material and Method

Triplicate data for each category were collected using following three microscopes at 100x magnification (Table 2.1).

i. First datasets were acquired using Labomed Digi 3 digital microscope (MS-1) which features an iVu 5100 digital camera module 5.0 megapixel CMOS sensor. The acquired images were 800x600 pixels with bit depth and resolution of 24 (eight per channel).
ii. Second datasets were acquired using a Motic BA210 digital microscope (MS-2) which features a Moticam 2500 digital camera module 5.0 megapixel CMOS sensor. The acquired images were 1280x1024 and 2592x1944 pixels with bit depth and resolution of 24 and 96 DPI respectively.
iii. Third datasets were acquired using Olympus CH20i digital microscope (MS-3) which features Smartphone digital camera with 16-megapixel BSI-CMOS sensor. The acquired images were 5312x2988 (16MP), 3984x2988(12MP) and 2048x1152(2.4MP) pixels with bit depth and resolution of 24 and 72 DPI respectively.

2.2.1 Description of Datasets

The architecture of Ziehl-Neelsen Sputum smears Microscopy image Database (ZNSM-iDB) with its applications is presented in Fig. 2.1. This database contains seven different categories of digital images in triplicate sets (one set from each microscope), which may be separately visualized or retrieved for further processing and applications (Table 2.1).

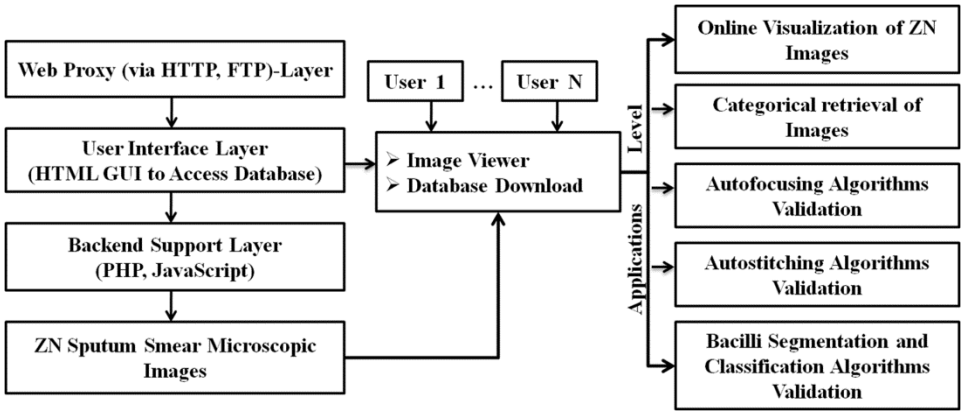

Fig. 2.1 Architecture and applications of ZNSM-iDB

Table 2. Category-wise presentation of datasets available in ZNSM-iDB

Group	Category of Data	No. of digital images from different Microscope (MS)		
		MS-1	MS-2	MS-3
1.	Autofocusing dataset*	9 stacks	10 stacks	30 stacks
2.	Overlapping view-fields for autostitching	7 Sets (50 to 90 Images/Set)	6 Sets (50 Images/ Set)	10 Sets (50 Images/Set)
3.	Manually segmented bacilli in a view-field	2 Sets (50 Images/Set)	2 Sets (50 Images/Set)	2 Sets (50 Images/Set)
4.	View-fields without bacilli	50	50	50
5.	Single or few bacilli	100	100	100
6.	Overlapping (occluded) bacilli	200	200	200
7.	Over-stained view-fields with bacilli and artifacts	250	250	250

* Each stack contains 20 images.

2.3 Results and Discussion

2.3.1 Application

The ZN sputum smear microscopy image database (ZNSM-iDB) is a unified image resource which has potential to assist the development of novel algorithms and tools related to automated grading (computer aided detection of bacilli) of smear slide using image processing techniques (Fig. 2.2). The database is freely available at http://14.139.240.55/znsm, and a tutorial is provided for effective exploration of diverse datasets.

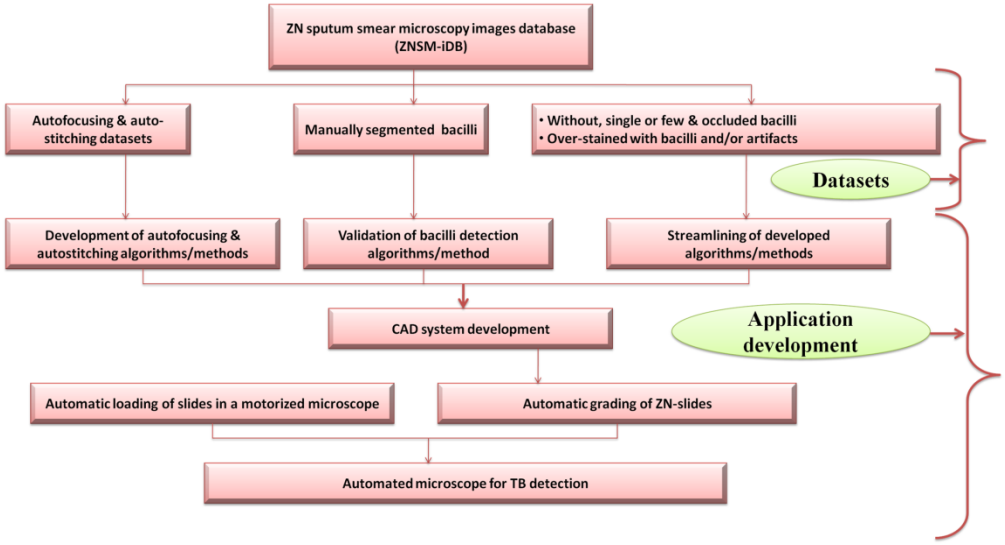
ZN sputum smear microscopy images database (ZNSM-iDB)
Autofocusing & auto-stitching datasets
Manually segmented bacilli
• Without, single or few & occluded bacilli
• Over-stained with bacilli and/or artifacts
Datasets
Development of autofocusing & autostitching algorithms/methods
Validation of bacilli detection algorithms/method
Streamlining of developed algorithms/methods
CAD system development
Application development
Automatic loading of slides in a motorized microscope
Automatic grading of ZN-slides
Automated microscope for TB detection

Fig. 2.2 Application of ZNSM-iDB database in automated microscopy

2.3.2 Validation

Autofocusing, autostitching, and bacilli segmentation and classification algorithms were implemented on ZNSM-iDB datasets. Comprehensive and comparative analyses of twenty-four focus measure functions (FMFs) have shown that Gaussian derivative, Tenengrad gradient, steerable filters, and Hemli and Scherer's mean are the most robust and accurate FMFs (refer Chapter 4). Different autostitching methods were evaluated on ZNSM-iDB datasets acquired using MS-1. The SIFT feature extraction and RANSAC feature selection methods were used to generate mosaic [35]. Similarity scores between the stitched and original images were 0.997, 0.988 and 0.989 using COR, SSIM and FSIM measures, respectively. These results were significantly better than the performances of Autostitch [36] and MicroMos [37] software (Fig.

2.3). Bacilli segmentation and classification was performed using watershed segmentation method on ZNSM-iDB datasets (refer to the Chapter 6 for detailed results).

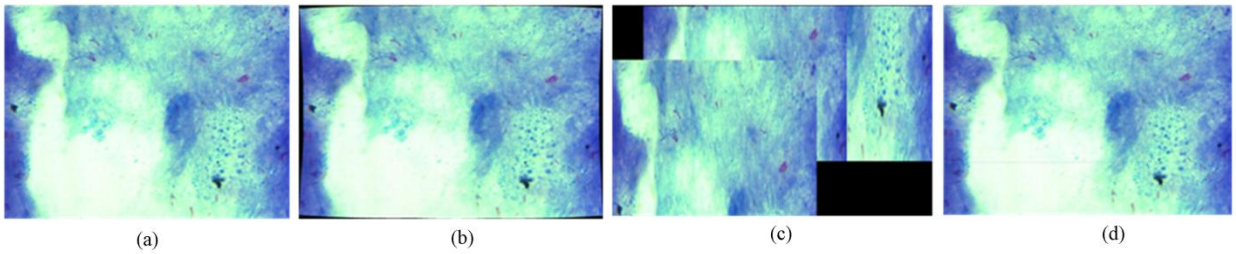

Fig. 2.3 Comparison between images of original view field with stitched mosaic. (a) Original image, (b) mosaic formed using Autostitch, (c) MicroMos and (d) divide-and-conquer.

2.4 Conclusion

The ZNSM-iDB is well diverse in terms of image acquisition technology and content. These images were acquired from three microscopes with different configurations and scope. Since autofocusing, autostitching and automated grading methods are necessary for automated microscope development; this resource may work as a standard platform to compare existing methods as well as to develop new algorithms. Diverse data available in this resource can facilitate the development of algorithms/model on one microscope, while the validation of the same method can be done on another microscope(s). The ZNSM-iDB is expected to serve as a referral resource for the development and validation of robust segmentation algorithms as the images possess high noise and artifacts.

CHAPTER # 3

IDENTIFICATION OF ROBUST FOCUS MEASURE FUNCTIONS FOR THE AUTOMATED CAPTURING OF FOCUSED IMAGES FROM ZIEHL-NEELSEN STAINED SPUTUM SMEAR MICROSCOPY SLIDE

Capturing of the best-focused image using focus measure function (FMF) from a stack of images acquired at different focus distances is a crucial step in automatic microscopy development. Detection of bacilli present in Ziehl-Neelsen (ZN) stained sputum smears conventional microscope (CM) images is critical to disease diagnosis. Studies have revealed that the performances of FMFs are sensitive to image contents and background noise. In this work, twenty-four diverse FMFs were implemented on thirty-one stacks of CM's field of view images acquired from three different microscopes to determine the best-focused one. Pre-processing techniques such as filtering and image distortions (noise, contrast, saturation, illumination, etc.) were performed to evaluate the robustness of every FMF. Gaussian derivative, steerable filters, Tenengrad, and Hemli and Scherer's FMFs were identified as the most robust and accurate functions with the accuracy >90%. These FMFs have a relatively less focus error and false maxima rate. These four FMFs can be implemented for automated capturing of the image from ZN-stained sputum smear slide.

3.1 Background

Higher sensitivity of popular, cost-effective and non-invasive smear microscopy (CM) test can bring better health outcomes against tuberculosis. The majority of current issues on tuberculosis diagnosis can be addressed by incorporating automated methods that increase the sensitivity and specificity of CM images by analyzing a large number of fields of view for *Mycobacterium tuberculosis* identification [26]. Automation prevents physical and mental fatigue, and reduces the time of field of view analysis [13]. Autofocusing, auto-stitching, and image segmentation and classification are the three sequential steps in automated microscopy system for the tuberculosis screening. Image focusing is very crucial in any automatic microscopy system as defocused images can affect the results of auto-stitching, and segmentation and classification [13]. Studies have also indicated that some FMFs performed significantly better in fluorescence microscopy

images [16-18] but produce an average outcome in CM. Although many FMFs were proposed for auto-focusing, their accuracies were mainly dependent on the content of images and instrument [8, 13, 19, 38, 39]. The effectiveness of FMFs also depends on the background of an image which are broadly grouped as the higher, medium and lower density [19, 40]. The CM images have higher density background than FM. Furthermore, improper Ziehl-Neelsen (ZN) staining dye leaves patches of blue regions which make focusing more difficult as focus functions consider these regions as significant features for identifying the best-focused one. Images acquired from ZN smear slide using CM are also distinctive due to their size, shape, debris and artifacts which make autofocusing even more difficult [13]. In the present study, twenty-four focus measure functions (FMFs) were selected on the basis of their performance in different auto-focusing applications. These FMFs were implemented on well-versed datasets from three different microscopes to identify efficient autofocus algorithms for CM images. These FMFs belong to five different categories according to their working principle. The change in performances of FMFs was also studied after inclusion of median filtering and image distortion using noise addition, contrast reduction, saturation increment and uneven illumination.

3.2 Material and Method

Total thirty-one autofocusing datasets (stacks) were retrieved from Ziehl-Neelsen Sputum smear Microscopy image Database (ZNSM-iDB) (Chapter 3).

3.2.1 Focus Measure Functions

The FMFs included in this study were implemented in MATLAB and broadly grouped into five major categories.

i. Gradient-based FMFs: These functions assume that well-focused image has more high-frequency content. Therefore, these FMFs use the gradient or first order derivative of the image to find well-focused one. Brenner gradient (BGR) [41], Energy gradient (EGR) [42], Gaussian derivative (GDR) [43], Thresholded absolute gradient (THR) [39], Squared gradient (SGR) [17], Tenengrad (TGR) [17, 44] & Tenengrad variance(VGR) [44] were used from this category.

ii. Laplacian-based FMFs: These functions also measure the amount of sharp edges in images but use the second-order of derivatives or Laplacian. Energy of Laplacian (ELP) [19], Modified Laplacian (MLP) [45] & Diagonal Laplacian (DLP) [46] were used from this category.

iii. Wavelet-based FMFs: These functions calculate the frequency and spatial contents of image using discrete wavelet transform coefficients. Sum of wavelet coefficients (SWC) [47], Variance of wavelet coefficients (VWC) [47] & Ratio of the wavelet coefficient (RWC) [48] were used from this category.

iv. Statistics-based FMFs: Functions in this group use various image statistics such as standard deviation, variance, autocorrelation, etc., to calculate the degree of focus of an image. Gray-level variance (GLV) [49], Gray-level local variance (LLV) [44], Normalized gray-level variance (GNV) [49], Standard deviation (SD) [15], Vollath's autocorrelation (VCR) [12, 16], Histogram entropy (EHS) [50] & Histogram Range (RHS) [50] were used from this category.

v. Other FMFs: Functions in this group are those which are not in above categories due to their working principles. Image curvature (ICR) [51], Hemli and Scherer's mean (HSM) [51], Steerable filters-base (SFB) [52] & Spatial frequency measure (SFM) [53] were used from this category.

3.2.2 Image Preprocessing

Median filtering, Poisson noise, contrast reduction, saturation and uneven illumination were incorporated to see the effect of different imaging conditions on focus measure outcomes.

3.2.3 Evaluation of focus measures

i. Accuracy criterion: The accuracy score of 1, 0.5 or 0 was assigned if the best-focused image from the stack was correctly classified; if the second best-focused was classified as the best-focused when the difference between the best and second-best image is marginal; or if the stack classification was failed, respectively. The accuracy in percent is calculated using Eq. 3.1:

$$\frac{\text{Sum of all score}}{\text{Total number of stacks}} \times 100 \qquad \textbf{(3.)}$$

The higher is the score, the more accurate is the algorithm.

ii. Focus error: The motor step difference between manually obtained and FMF predicted best focus image was considered as the focus error. Each motor step represents a different focus distance. The difference between two adjacent motor steps is 2-3 µm.

iii. Number of false maxima: The number of false maxima produced by an FMF was determined. It is the total number of maxima present in a focus curve excluding the global maxima.

iv. Full width at half maximum (FWHM): This criterion was used to check the more appropriate FMF for optimization process when the performances of focus functions are similar.

3.3 Results and Discussion

Average performances of twenty FMFs were computed with and without preprocessing on overall datasets from three microscopes. The Gaussian derivative (GDR), Tenengrad (TGR), Helmli's measure (HSM), thresholded gradient (THR) and steerable filter-based (SFB) focus measures were ranked as top five FMFs based on their accuracy (>90), focus error and number of false maxima (Table 3.1, Fig. 3.1). THR performances were affected by different imaging conditions; therefore, it was not considered as robust FMF. Mean full width at half maximum (FWHM) of top four focus measure functions: GDR = (2.68 (before DC removal), 2.53 (after DC Removal)), TGR = (5.35, 3.30), HSM = (6.72, 3.35) and SFB = (3.52, 3.27). Lesser DC offset value indicates the significance of an FMF. Lesser the FWHM value, better is the focus measure function.

Table 3.1 Performance of focus measures functions (FMFs) without pre-processing

Methods*	Accuracy	Focus Error	False Max	Overall Score	Overall Rank
GDR	96.77(1)	0.03(1)	0.07(1)	3	1
TGR	96.77(1)	0.05(2)	0.1(2)	5	2
HSM	95.16(2)	0.05(2)	0.28(5)	9	3
THR	95.16(2)	0.34(6)	0.1(2)	10	4
SFB	90.32(3)	0.39(7)	0.07(1)	11	5
GNV	88.71(4)	0.27(5)	0.17(3)	12	6
GLV	87.1(5)	0.16(3)	0.21(4)	12	6
ICR	95.16(2)	0.03(1)	0.69(11)	14	7
SD	88.71(4)	0.6(9)	0.17(3)	16	8
SFM	90.32(3)	0.39(7)	0.41(9)	19	9
VCR	85.48(6)	0.66(10)	0.21(4)	20	10
VGR	80.65(7)	0.23(4)	0.41(9)	20	10
EGR	75.81(8)	1.15(12)	0.34(7)	27	11
SGR	74.19(9)	0.52(8)	0.48(10)	27	11
MLP	75.81(8)	1.37(13)	0.34(7)	28	12
EHS	74.19(9)	0.68(11)	0.38(8)	28	12
DLP	72.58(10)	1.92(15)	0.31(6)	31	13
SWC	64.52(12)	1.94(16)	0.17(3)	31	13

ELP	67.74(11)	2.1(17)	0.76(12)	40	14
VWC	62.9(13)	1.66(14)	6.03(13)	40	14

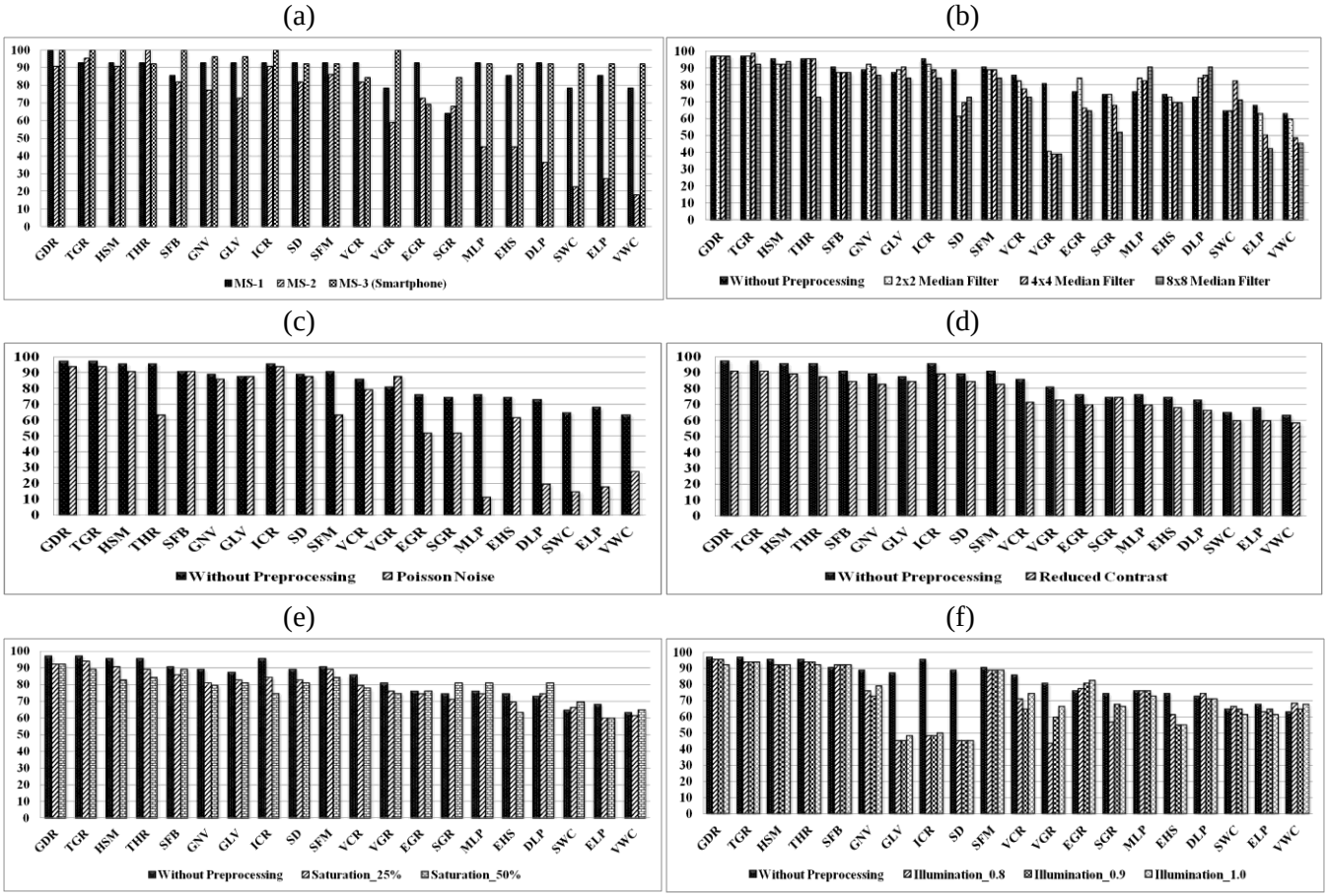
(a)
(b)
(c)
(d)
(e)
(f)
GDR TGR HSM THR SFB GNV GLV ICR SD SFM VCR VGR EGR SGR MLP EHS DLP SWC ELP VWC
MS-1 MS-2 MS-3 (Smartphone)
Without Preprocessing 2x2 Median Filter 4x4 Median Filter 8x8 Median Filter
Without Preprocessing Poisson Noise
Without Preprocessing Reduced Contrast
Without Preprocessing Saturation_25% Saturation_50%
Without Preprocessing Illumination_0.8 Illumination_0.9 Illumination_1.0

Fig. 3.1 Microscopes-wise accuracy and mean accuracy in percentage of FMFs. (a) Without pre-processing, (b) median filtering, (c) noise addition, (d) contrast reduction, (e) increase saturation level & (f) uneven illumination.

3.4 Conclusion

The identification of efficient FMFs for a particular imaging system assumes lots of significance for the development of auto-focusing instrument. A comprehensive analyses of twenty FMFs on diverse bright-field CM images from three different microscopes provided GDR, TGR, SFB and HSM as the most robust and accurate FMFs as they showed the best overall accuracy, and performances were invariant to pre-processing and image distortions. These FMFs can be implemented for the development of new automated microscopy system which can capture an image directly from ZN stained sputum smear slide.

CHAPTER # 4

ESTABLISHMENT OF HYBRIDIZED FOCUS MEASURE FUNCTIONS AS UNIVERSAL METHOD FOR AUTO-FOCUSING

Exact focusing is essential for any automatic image capturing system. Performances of focus measure functions (FMFs) used for autofocusing are sensitive to image contents and imaging systems. Therefore, identification of universal FMF assumes a lot of significance. In this study, eight FMFs (performed effectively in ZN and FM image modalities) were hybridized in a pair of two and implemented simultaneously on a single stack to identify the best focus point. In total, 28 hybrid FMFs (HFMFs) and eight FMFs were implemented on stacks of images from three different imaging modalities. Performance of FMFs was found to the best at 50% region-sampling. Accuracy, focus error and false maximum were calculated to evaluate the performance of each FMF. Nineteen HFMFs provided more than 90% accuracy. Image distortion (noise, contrast, saturation, illumination, etc.) were performed to evaluate robustness of HFMFs. Hybrid of Tenengrad variance and steerable filter based (VGRnSFB) FMFs was identified as the most robust and accurate function with the accuracy of ≥90% and relatively less focus error and false maximum rate. Sharpness of the focus curve of VGRnSFB along with eight individual FMFs was also computed for determining the efficacy of HFMF for optimization process. VGRnSFB HFMF can be implemented for automated capturing of an image for any imaging system.

4.1 Material and Method

Three different image modalities containing 87 stacks of images were used to evaluate the performance of individual and hybrid focus measure functions. *ZN Sputum Smear Microscopy (31 stacks,* Fig. 4.1, Chapter 2), *Fluorescent Sputum Smear Microscopy (35 stacks, Fig. 4.1,* [14]) and *Multi-Spectral (MS) Dataset (21 stacks,* Fig. 4.1, [54]*)* with visible, near-infrared and thermal spectrum were used in this study.

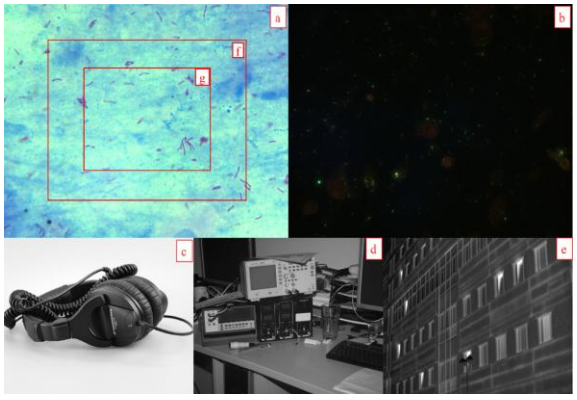

Fig. 4.1 Image modalities used to evaluate HFMFs. (a) Image acquired from ZN sputum smear conventional microscopy (CM), (b) Image acquired from sputum smear fluorescent microscopy (FM), (c) Image acquired in visible spectrum (VS), (d) Image acquired in near-infrared spectrum, (e) Image acquired in thermal spectrum (TS), (f) Depiction of the 50% region-sampled area, and (g) Depiction of the 25% region-sampled area used to evaluate focus measure functions

4.1.1 Focus Measure Functions (FMFs)

Eight most common FMFs were included in this study as their performances in ZN (CM) and FM higher density images were good (Chapter 3) [13, 14,]. These three major categories FMFs and their HFMFs were implemented in MATLAB (Table 4.1).

Table 4.1 Focus measure functions (FMFs) and their category used to form hybrid-FMF for identifying the best-focused images.

S. No.	Category	Focus Measure Function (FMF)	Reference
1.	Gradient-based	Gaussian derivative (GDR)	[43]
2.		Tenengrad (TGR)	[18]
3.		Tenengrad variance(VGR)	[44]
4.	Statistics-based	Normalized gray-level variance (GNV)	[49]
5.		Vollath's autocorrelation (VCR)	[17, 18]
6.	Other	Hemli and Scherer's mean (HELM)	[51]
7.		Steerable filters-based (SFB)	[52]
8.		Spatial frequency measure (SFM)	[53]

4.1.2 Hybridization

Hybridization of FMFs means two FMFs are implemented simultaneously on a single stack to calculate the hybrid focus measure using Eq. 4.1.

$$Hybrid\ FMF\ (HFMF) = \frac{FMF1 + FMF2}{2} \qquad \textbf{(4.1)}$$

Where, FMF1 and FMF2 are two focus measure functions which were implemented simultaneously.

4.1.3 Region-sampling& Preprocessing

Region-sampling was performed to implement FMFs on 25%, 50% and 75% parts of whole image (Fig. 4.1f, 4.1g). Region-sampling was performed to achieve better accuracy as well as to reduce computation time [40]. Image Preprocessing such as Poisson noise, saturation, contrast and uneven illumination was incorporated in images to check the robustness of HFMFs in different imaging conditions.

4.1.4 Evaluation of focus measures

Accuracy, Focus error and false maxima were calculated to evaluate the hybrid FMFs (refer Chapter 4). Finally, *Sharpness of focus curve* is used to identify the FMF and HFMFs with better convergence rate to the best focus point.

4.2 Results and Discussion

This study proposed an autofocus system using hybrid focus measure function (HFMF) and assumed that some HFMFs could be the robust across diverse image modalities as well as different imaging conditions (noise, saturation, etc.). HFMFs analyses were performed only on 50% region-sampled images as the result was optimal and mean computation time was minimal at this level (Table 4.2). Improved performance of FMFs on 50% and 75% region-sampling might be due to better focusing on central part of the image than boundaries. Performances of FMFs and HFMFs in different imaging conditions are presented in Fig. 4.2. VGRnSFB HFMF has been identified as the most robust and accurate autofocus function as it is invariant to different imaging conditions with higher accuracy rate, and minimum focus error and false maximum. VGRnSFB has also produced a comparable sharpness curve (Fig. 4.3), and it can be implemented in real time system.

Table 4.2 Accuracy of focus measure functions (FMFs) in percent with different region-sampling data *i.e.* 25%, 50%, 75% and whole image

Method	25%	50%	75%	Whole Image
GDR	90.1	94.8	92.4	92.4
GNV	46.5	46.5	46.5	47.1
HELM	80.2	82.6	79.1	79.1
SFB	54.7	60.5	56.4	58.1
SFM	86.6	88.4	89.0	88.4
TGR	90.7	95.3	93.6	95.9
VCR	76.2	82.6	84.9	87.2
VGR	84.9	87.8	89.5	85.5

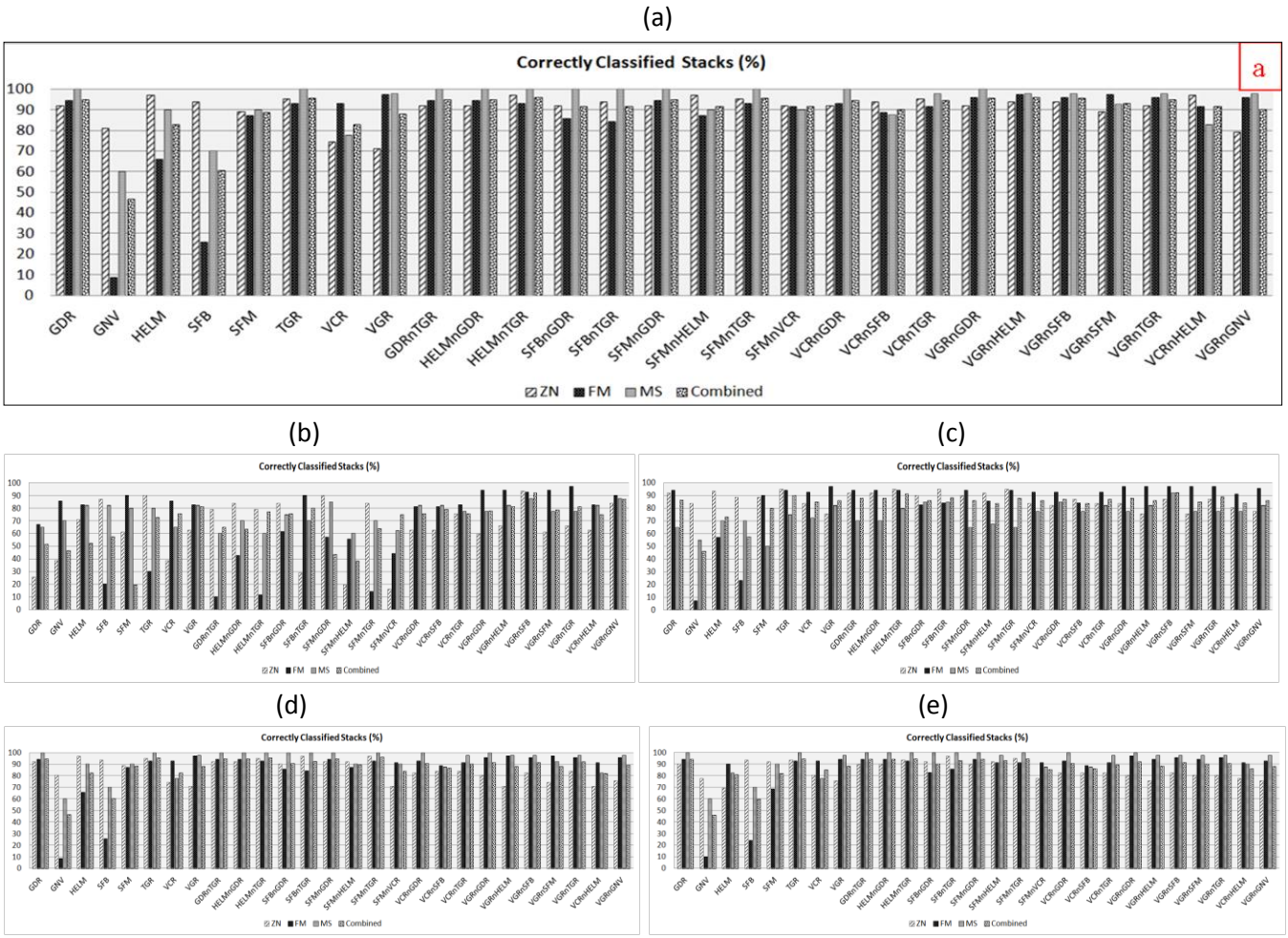

Fig. 4.2 Accuracies of FMFs and HFMFs in percent at 50% region-sampling data. (a) Without preprocessing, (b) noise addition, (c) saturation, (d) contrast reduction, & (e) uneven illumination.

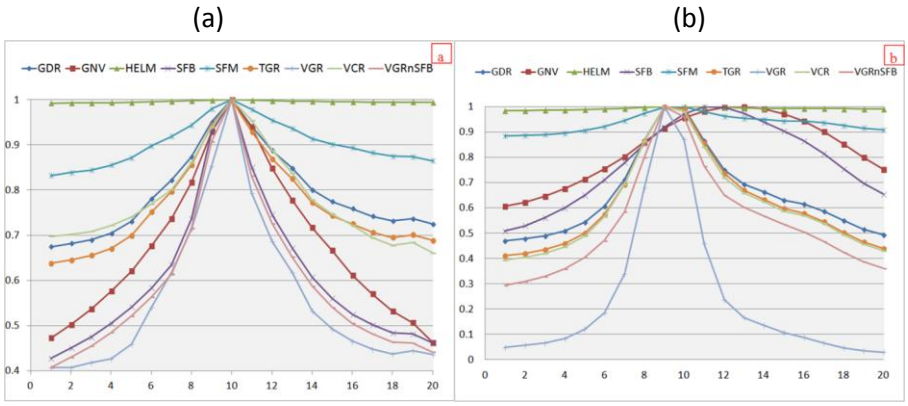

Fig. 4.3 Sharpness curve of nine focus measure functions (8 FMF and one HFMF) including hybrid FMF (VGRnSFB). Narrow peak represents rapid convergence of FMF.

4.3 Conclusion

Exact autofocusing using FMFs is very crucial step of any imaging system. Studies have reported that the performance of FMFs is sensitive to image contents. A comprehensive analysis of twenty-eight hybrid FMFs on diverse datasets, spanning almost all image categories, provided nineteen hybrid methods with accuracy >90%. Effectiveness of these HFMFs was tested under different imaging conditions. VGRnSFB is found to be the most robust and accurate HFMF as it showed the best overall accuracy and robustness as the performances were independent to different image distortions. This HFMF may be implemented in any imaging system which can capture the best focused image automatically.

CHAPTER # 5

AUTOMATIC DETECTION AND CLASSIFICATION OF TUBERCULOSIS BACILLI FROM ZN-STAINED SPUTUM SMEAR IMAGES USING WATERSHED SEGMENTATION

In this chapter, an automated method for detection and classification of bacilli using watershed segmentation is presented. The efficacy of this algorithm to distinguish overlapping bacilli was also discussed. Several preprocessing techniques have been implemented prior to watershed segmentation for removal of the artifacts and objects larger or smaller than bacilli. Segmentation and classification of bacilli was performed using datasets from ZNSM-iDB database.

5.1 Material and Method

Forty images from three microscopes (MS-1, MS-2 and MS-3) were randomly extracted and grouped into medium- and high-density background datasets. Shape and size of the objects were determined to filter the true bacilli. Similarly, the watershed algorithm was implemented on randomly extracted thirty images from Smartphone enabled microscope (MS-3) to segment bacilli. Additionally, images were divided into four groups based on the infection level (Table 5.1) [56], and sensitivity and specificity of watershed segmentation method for classifying an image as TB positive or negative. Sensitivity and precision rate of this segmentation method for identifying true bacilli were also determined for each group. Furthermore, discordance rate was calculated for watershed segmentation method to evaluate the percentage of pairs where the observation with TB-positive has a lower predicted probability than TB-negative [56, 57]. Predicted probability was calculated in binary logistic regression model [56].

Table 5.1 Grading of view-fields on the basis of infection level.

Number of Bacilli	Number of View-Fields to be Examined	Grading
1-9 in 100 view-fields	100	Scanty[a]
10-99 in 100 view-fields	100	1+
1-10 in each view-field	50	2+
>10 in each view-field	20	3+

[a]Report exact number of bacilli present in the view-fields

5.1.1 Method

Bacilli detection and classification have been performed in six different steps which are presented as flowchart in Fig. 5.1. In first step, the collected images are converted to greyscale for further processing and implementation of watershed segmentation. Second step is used to enhance the contrast of greyscale image for better and clear visibility of bacilli. Binary conversion of image through thresholding is performed in third step to separate bacilli from the background. In step four, only bacilli are filtered out by removing objects which are smaller or larger than bacilli and not in rod shape. Pixels of connected components (objects) have been used for calculating the size of object while the perimeter of objects is used in Eq. 5.1 for calculating the circularity of objects. In final step, the watershed algorithm is implemented for segmentation and separation of overlapping (touching) bacilli.

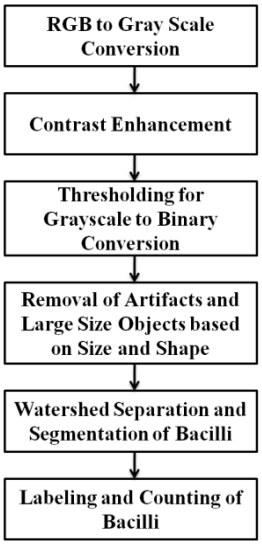
RGB to Gray Scale Conversion
Contrast Enhancement
Thresholding for Grayscale to Binary Conversion
Removal of Artifacts and Large Size Objects based on Size and Shape
Watershed Separation and Segmentation of Bacilli
Labeling and Counting of Bacilli

Fig. 5.1 Flowchart of proposed method

$$\textbf{Circularity} = \frac{4*\pi*\textbf{area}}{\textbf{Perimeter}^2} \qquad (5.1)$$

5.1.2 Watershed Transform

The watershed transform is a popular segmentation method of mathematical morphology field. Beucher and Lantue have introduced it first time in image analysis [58]. The basic concept of watershed is based on visualizing a gray level image into its topographic representation (Fig. 5.2), where the height of every pixel is directly related to its gray level, and considers rain gradually falling on the terrain or land relief, and then the watersheds lines separate the

catchment basins. In order to locate catchment basin boundaries at high gradient points, gradient of the image is generally used to compute the watershed transform [59].

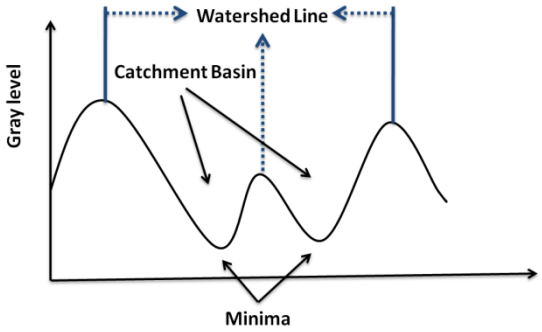

Fig. 5.2 Topographical representation of a gray level image using Watershed method.

5.2 Results and Discussion

Fig. 5.3 (a-f) shows that proposed segmentation technique is able to extract most of the bacilli and eliminate unwanted objects from ZN stained images. The sensitivity and specificity of this method for classifying a medium density background image as tuberculosis positive or negative were 100% and 93%, respectively, while for a high-density background sensitivity remained unchanged, but specificity was reduced to 72% due to over-staining and artifacts. Similarly, the sensitivity and specificity of this segmentation method for classifying a Smartphone enabled microscopic images (medium to high-density background) as tuberculosis bacilli positive or negative were 93.3% and 87%, respectively. The sensitivity, specificity and Precision rate of watershed segmentation on different infection level (Scanty, 1+, 2+ and 3+) are depicted in Table 5.3. Furthermore, this segmentation provided a discordant rate of 3.73% with respect to tuberculosis positive and negative cases.

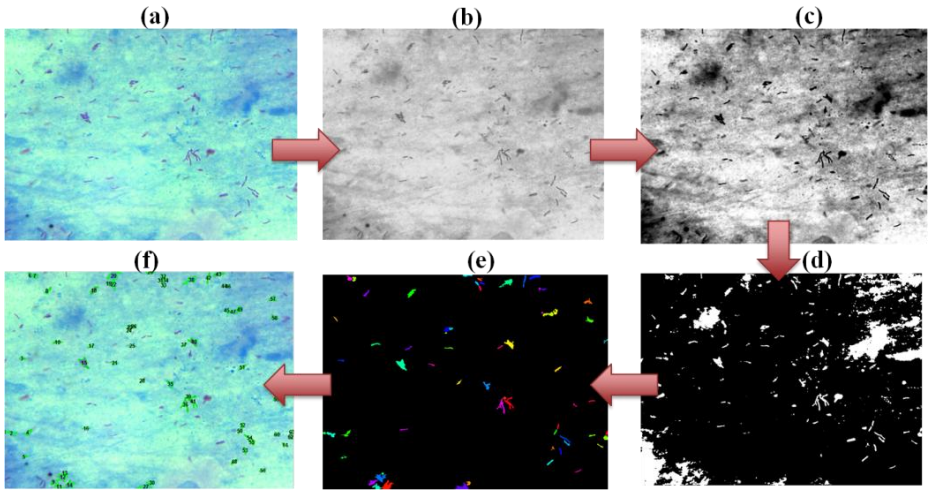

Fig. 5.3 Stepwise presentation of Watershed segmentation method for detection and classification of bacilli in ZN stained sputum smear image. (a) Original image, (b) RGB to grayscale conversion, (c) Contrast enhancement, (d) Binarized image, (e) Artifacts removal and watershed segmentation & (f) Bacilli labeled Image

Table 5.3 Performance of watershed segmentation method on different infection level.

Grade	Classification of View-fields as TB Positive or Negative (%)		Bacilli Segmentation in a View-field (%)	
	Sensitivity[a]	Specificity[a]	Sensitivity[b]	Precision[b]
Scanty	42.86	74.73	55.56	26.19
1+	81.25	75.29	55.56	52.4
2+	92.86	NA[*]	61.73	72
3+	100	NA	90.23	79.09

[a]Sensitivity and specificity of watershed segmentation method for classifying an image (view-field) as TB positive or negative.

[b]Sensitivity and precision rate of watershed segmentation method for the segmentation of bacilli in the view-field images.

[*]Datasets does not contain tuberculosis negative view-field images.

5.3 Conclusion

The performance of this segmentation method is related to the level of infection where higher level of infection yields better performance. Method is very significant in detecting bacilli even for challenging images with unclear and overlapped bacilli. Cost effective and computationally powerful camera-enable Smartphone microscopy could be used as an alternative of higher cost and maintenance digital microscopes. It will not only work as a digital microscope but it would provide access to EHR, automatic analysis of samples, better diagnostic and monitoring.

CHAPTER # 6

CONCLUSION AND FUTURE WORK

6.1 Conclusion

In the proposed thesis, work in support of automated microscopy development using image analysis techniques has been performed. Current work advances the knowledge of automated microscopy development by providing autofocusing, and bacilli segmentation methods. Automated microscope developed through these methods may provide the mass screening of cases in TB-endemic countries. Automation will also deliver faster results which will assist in early diagnosis of tuberculosis. It will also increase the sensitivity and specificity of tuberculosis diagnosis using ZN sputum smear test. Researchers can validate their automated microscopy methods on unified database, "ZNSM-iDB".

6.2 Future Work

In the future, sensitivity and specificity of watershed segmentation method may be optimized for better disease diagnosis. Following improvements can be incorporated in existing approach:

- Use of color feature in bacilli segmentation to remove overstrained blue artifacts and retain bacilli with reddish color.
- Incorporation of machine learning for supervised classification of TB positive and negative images.
- Comparative analysis of watershed algorithm with previously reported techniques.

REFERENCES

1. Shortliffe, E.H. and Cimino, J.J. eds., 2013. *Biomedical informatics: computer applications in health care and biomedicine*. Springer Science & Business Media.
2. McNerney, R., 2015. Diagnostics for developing countries. *Diagnostics*, *5*(2), pp.200-209.
3. Small, P. M., & Pai, M., 2010. Tuberculosis diagnosis—time for a game change. *N. Engl. J. Med*, 363(11), 1070-1071.
4. Tadrous, P.J., 2010. Computer-Assisted Screening of Ziehl-Neelsen–Stained Tissue for Mycobacteria: Algorithm Design and Preliminary Studies on 2,000 Images. *American journal of clinical pathology*, *133*(6), pp.849-858.
5. Khutlang, R., Krishnan, S., Whitelaw, A. and Douglas, T.S., 2010. Automated detection of tuberculosis in Ziehl-Neelsen-stained sputum smears using two one-class classifiers. *Journal of microscopy*, *237*(1), pp.96-102.
6. Ba, F. and Rieder, H.L., 1999. A comparison of fluorescence microscopy with the Ziehl-Neelsen technique in the examination of sputum for acid-fast bacilli. *The International Journal of Tuberculosis and Lung Disease*, *3*(12), pp.1101-1105.
7. Sotaquira, M., Rueda, L. and Narvaez, R., 2009, March. Detection and quantification of bacilli and clusters present in sputum smear samples: a novel algorithm for pulmonary tuberculosis diagnosis. In *Digital Image Processing, 2009 International Conference on* (pp. 117-121). IEEE.
8. Junior, A.K., Costa, M.G., Costa Filho, C.F., Fujimoto, L.B. and Salem, J., 2010, August. Evaluation of autofocus functions of conventional sputum smear microscopy for tuberculosis. In *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE* (pp. 3041-3044). IEEE.
9. Shingadia, D. and Novelli, V., 2003. Diagnosis and treatment of tuberculosis in children. *The Lancet infectious diseases*, *3*(10), pp.624-632.
10. Mesker, W.E., Torrenga, H., Sloos, W.C.R., Vrolijk, H., Tollenaar, R.A.E.M., de Bruin, P.C., van Diest, P.J. and Tanke, H.J., 2004. Supervised automated microscopy increases sensitivity and efficiency of detection of sentinel node micrometastases in patients with breast cancer. *Journal of clinical pathology*, *57*(9), pp.960-964.
11. Costa, M.G.F., Costa Filho, C.F.F., Kimura, A., Levy, P.C., Xavier, C.M. and Fujimoto, L.B., 2014, August. A sputum smear microscopy image database for automatic bacilli detection in conventional microscopy. In *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE* (pp. 2841-2844). IEEE.
12. Sawadogo, T.L., Savadogo, L.G., Diande, S., Ouedraogo, F., Mourfou, A., Gueye, A., Sawadogo, I., Nebié, B., Sangare, L. and Ouattara, A.S., 2012. Comparison of Kinyoun,

auramine O, and Ziehl-Neelsen staining for diagnosing tuberculosis at the National Tuberculosis Center in Burkina Faso. *Medecine et sante tropicales*, *22*(3), pp.302-306.

13. Osibote, O.A., Dendere, R., Krishnan, S. and Douglas, T.S., 2010. Automated focusing in bright-field microscopy for tuberculosis detection. *Journal of microscopy*, *240*(2), pp.155-163.
14. Mateos-Pérez, J.M., Redondo, R., Nava, R., Valdiviezo, J.C., Cristóbal, G., Escalante-Ramírez, B., Ruiz-Serrano, M.J., Pascau, J. and Desco, M., 2012. Comparative evaluation of autofocus algorithms for a real-time system for automatic detection of Mycobacterium tuberculosis. *Cytometry Part A*, *81*(3), pp.213-221.
15. Costa Filho, C.F., Costa, M.F. and Júnior, A.K., 2012. Autofocus functions for tuberculosis diagnosis with conventional sputum smear microscopy. *Proc. of Current Microscopy Contributions to Advances in Science and Technology. Formatex Research Center*, pp.13-20.
16. Vollath, D., 1987. Automatic focusing by correlative methods. *Journal of Microscopy*, *147*(3), pp.279-288.
17. Vollath, D., 1988. The influence of the scene parameters and of noise on the behaviour of automatic focusing algorithms. *Journal of microscopy*, *151*(2), pp.133-146.
18. Santos, A., Ortiz de Solórzano, C., Vaquero, J.J., Pena, J.M., Malpica, N. and Del Pozo, F., 1997. Evaluation of autofocus functions in molecular cytogenetic analysis. *Journal of microscopy*, *188*(3), pp.264-272.
19. Russell, M.J. and Douglas, T.S., 2007, August. Evaluation of autofocus algorithms for tuberculosis microscopy. In *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE* (pp. 3489-3492). IEEE.
20. Yang, F., Deng, Z.S. and Fan, Q.H., 2013. A method for fast automated microscope image stitching. *Micron*, *48*, pp.17-25.
21. Costa Filho, C.F.F., Levy, P.C., Xavier, C.D.M., Fujimoto, L.B.M. and Costa, M.G.F., 2015. Automatic identification of tuberculosis mycobacterium. *Research on Biomedical Engineering*, *31*(1), pp.33-43.
22. Priya, E. and Srinivasan, S., 2015. Automated Identification of Tuberculosis Objects in Digital Images Using Neural Network and Neuro Fuzzy Inference Systems. *Journal of Medical Imaging and Health Informatics*, *5*(3), pp.506-512.
23. Veropoulos, K., Campbell, C. and Learmonth, G., 1998. Image processing and neural computing used in the diagnosis of tuberculosis. 8-8.
24. de Jager, K., Fickling, S., Krishnan, S., Jabbari, M., Learmonth, G.W. and Douglas, T.S., 2014. Automated fluorescence microscope for tuberculosis detection. *Journal of Medical Devices*, *8*(3), p.030943.

25. Priya, E. and Srinivasan, S., 2013. Automated decision support system for tuberculosis digital images using evolutionary learning machines. *European Journal for Biomedical Informatics*, *9*(2), pp.3-8.
26. Tapley, A., Switz, N., Reber, C., Davis, J.L., Miller, C., Matovu, J.B., Worodria, W., Huang, L., Fletcher, D.A. and Cattamanchi, A., 2013. Mobile digital fluorescence microscopy for diagnosis of tuberculosis. *Journal of clinical microbiology*, *51*(6), pp.1774-1778.
27. CostaFilho, C.F., Levy, P.C., Xavier, C.M., Costa, M.G., Fujimoto, L.B. and Salem, J., 2012, August. Mycobacterium tuberculosis recognition with conventional microscopy. In *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE* (pp. 6263-6268). IEEE.
28. Zhai, Y., Liu, Y., Zhou, D. and Liu, S., 2010, December. Automatic identification of mycobacterium tuberculosis from ZN-stained sputum smear: Algorithm and system design. In *Robotics and Biomimetics (ROBIO), 2010 IEEE International Conference on* (pp. 41-46). IEEE.
29. Osman, M.K., Mashor, M.Y. and Jaafar, H., 2011, March. Tuberculosis bacilli detection in Ziehl-Neelsen-stained tissue using affine moment invariants and Extreme Learning Machine. In *Signal Processing and its Applications (CSPA), 2011 IEEE 7th International Colloquium on* (pp. 232-236). IEEE.
30. Armato, S.G., McLennan, G., Bidaut, L., McNitt-Gray, M.F., Meyer, C.R., Reeves, A.P., Zhao, B., Aberle, D.R., Henschke, C.I., Hoffman, E.A. and Kazerooni, E.A., 2011. The lung image database consortium (LIDC) and image database resource initiative (IDRI): a completed reference database of lung nodules on CT scans. *Medical physics*, *38*(2), pp.915-931.
31. Depeursinge, A., Vargas, A., Platon, A., Geissbuhler, A., Poletti, P.A. and Müller, H., 2012. Building a reference multimedia database for interstitial lung diseases. *Computerized medical imaging and graphics*, *36*(3), pp.227-238.
32. Song, Y., Cai, W., Zhou, Y. and Feng, D.D., 2013. Feature-based image patch approximation for lung tissue classification. *IEEE transactions on medical imaging*, *32*(4), pp.797-808.
33. de Carvalho Filho, A.O., de Sampaio, W.B., Silva, A.C., de Paiva, A.C., Nunes, R.A. and Gattass, M., 2014. Automatic detection of solitary lung nodules using quality threshold clustering, genetic algorithm and diversity index. *Artificial intelligence in medicine*, *60*(3), pp.165-177.
34. Diciotti, S., Lombardo, S., Falchini, M., Picozzi, G. and Mascalchi, M., 2011. Automated segmentation refinement of small lung nodules in CT scans by local shape analysis. *IEEE Transactions on Biomedical Engineering*, *58*(12), pp.3418-3428.

35. Chauhan, A. Development of efficient computer aided diagnosis methods for tuberculosis diagnosis (Unpublished doctoral thesis). Jaypee University of Information Technology, India. (2016)
36. Brown, M. and Lowe, D.G., 2007. Automatic panoramic image stitching using invariant features. *International journal of computer vision*, *74*(1), pp.59-73.
37. Piccinini, F., Bevilacqua, A. and Lucarelli, E., 2013. Automated image mosaics by non-automated light microscopes: the MicroMos software tool. *Journal of microscopy*, *252*(3), pp.226-250.
38. Panicker, R.O., Soman, B., Saini, G. and Rajan, J., 2016. A review of automatic methods based on image processing techniques for tuberculosis detection from microscopic sputum smear images. *Journal of medical systems*, *40*(1), p.17.
39. Pertuz, S., Puig, D. and Garcia, M.A., 2013. Analysis of focus measure operators for shape-from-focus. *Pattern Recognition*, *46*(5), pp.1415-1432.
40. Gu, C.C., Wu, K.J., Hu, J., Hao, C. and Guan, X.P., 2015. Region sampling for robust and rapid autofocus in microscope. *Microscopy research and technique*, *78*(5), pp.382-390.
41. Brenner, J.F., Dew, B.S., Horton, J.B., King, T., Neurath, P.W. and Selles, W.D., 1976. An automated microscope for cytologic research a preliminary evaluation. *Journal of Histochemistry & Cytochemistry*, *24*(1), pp.100-111.
42. Subbarao, M., Choi, T. and Nikzad, A., 1993. Focusing techniques. *Optical engineering-bellingham-international society for optical engineering-*, *32*, pp.2824-2824.
43. Geusebroek, J.M., Cornelissen, F., Smeulders, A.W. and Geerts, H., 2000. Robust autofocusing in microscopy. *Cytometry Part A*, *39*(1), pp.1-9.
44. Pech-Pacheco, J.L., Cristóbal, G., Chamorro-Martinez, J. and Fernández-Valdivia, J., 2000. Diatom autofocusing in brightfield microscopy: a comparative study. In *Pattern Recognition, 2000. Proceedings. 15th International Conference on* (Vol. 3, pp. 314-317). IEEE.
45. Nayar, S.K. and Nakagawa, Y., 1994. Shape from focus. *IEEE Transactions on Pattern analysis and machine intelligence*, *16*(8), pp.824-831.
46. Thelen, A., Frey, S., Hirsch, S. and Hering, P., 2009. Improvements in shape-from-focus for holographic reconstructions with regard to focus operators, neighborhood-size, and height value interpolation. *IEEE Transactions on Image Processing*, *18*(1), pp.151-157.
47. Yang, G. and Nelson, B.J., 2003, October. Wavelet-based autofocusing and unsupervised segmentation of microscopic images. In *Intelligent Robots and Systems, 2003.(IROS 2003). Proceedings. 2003 IEEE/RSJ International Conference on* (Vol. 3, pp. 2143-2148). IEEE.
48. Xie, H., Rong, W. and Sun, L., 2006, October. Wavelet-based focus measure and 3-d surface reconstruction method for microscopy images. In *Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on* (pp. 229-234). IEEE.

49. Groen, F.C., Young, I.T. and Ligthart, G., 1985. A comparison of different focus functions for use in autofocus algorithms. *Cytometry*, *6*(2), pp.81-91.
50. Firestone, L., Cook, K., Culp, K., Talsania, N. and Preston, K., 1991. Comparison of autofocus methods for automated microscopy. *Cytometry*, *12*(3), pp.195-206.
51. Helmli, F.S. and Scherer, S., 2001. Adaptive shape from focus with an error estimation in light microscopy. In *Image and Signal Processing and Analysis, 2001. ISPA 2001. Proceedings of the 2nd International Symposium on* (pp. 188-193). IEEE.
52. Minhas, R., Mohammed, A.A. and Wu, Q.J., 2012. An efficient algorithm for focus measure computation in constant time. *IEEE Transactions on Circuits and Systems for Video Technology*, *22*(1), pp.152-156.
53. Huang, W. and Jing, Z., 2007. Evaluation of focus measures in multi-focus image fusion. *Pattern recognition letters*, *28*(4), pp.493-500.
54. Zukal, M., Mekyska, J., Cika, P. and Smekal, Z., 2013. Interest points as a focus measure in multi-spectral imaging. *Radioengineering*, *22*(1), pp.68-81.
55. Revised National Tuberculosis Control Programme. Module for laboratory technicians. New Delhi: Directorate General of Health Services, Ministry of Health and Family Welfare, 15. 1999.
56. Dayton, C.M., 1992. Logistic regression analysis. *Stat*, pp.474-574.
57. Bhalla, D, "Calculating Concordant, Discordant and Tied Pairs", (accessed 11 April 2017). [http://www.listendata.com/2014/08/modeling-tips-calculating-concordant.html]
58. Beucher, S., Lantuejoul, C., 1979. Use of watersheds in contour detection", In *International Workshop on Image Processing, Real-Time Edge and Motion Detection/ Estimation*, (1979).
59. Tarabalka, Y., Chanussot, J. and Benediktsson, J.A., 2010. Segmentation and classification of hyperspectral images using watershed transformation. *Pattern Recognition*, *43*(7), pp.2367-2379.

PUBLICATIONS

Journal Publication

1. **Shah MI,** Mishra S, Yadav VK , Chauhan A, Sarkar M, Sudarshan SK, Rout C. "Ziehl–Neelsen sputum smear microscopy image database: a resource to facilitate automated bacilli detection for tuberculosis diagnosis", *J. Med. Imag*. 4(2), 027503 (Jun 30, 2017).; **DOI: 10.1117/1.JMI.4.2.027503 (**IF: N/A, **Scopus Indexed,** SPIE**)**

2. **Shah MI,** Mishra S, Sarkar M, Rout C. "Identification of robust focus measure functions for the automated capturing of focused images from Ziehl–Neelsen stained sputum smear microscopy slide." Cytometry Part A. 2017 Jun 2. **DOI: 10.1002/cyto.a.23142 (IF: 3.22, Scopus Indexed,** Wiley**)**

3. **Shah MI,** Mishra S, Rout C. "Establishment of hybridized focus measure functions as universal method for auto-focusing." J Biomed Opt. SPIE (In Revision) (**IF: 2.53, Scopus Indexed,** SPIE)

Conference Publication

1. **Shah MI**, Mishra S, Sarkar M, Sudarshan SK. "Automatic detection and classification of tuberculosis bacilli from camera-enabled smartphone microscopic images." In Fourth International Conference on Parallel, Distributed and Grid Computing (PDGC), 2016 Dec (pp. 287-290). IEEE. DOI: 10.1109/PDGC.2016.7913161 **(**IF: N/A, **Scopus Indexed)**

2. **Shah MI**, Mishra S, Sarkar M, Rout C. "Automatic Detection and Classification of Tuberculosis Bacilli from ZN-stained Sputum Smear Images using Watershed Segmentation." In International Conference on Signal Processing (ICSP), 2016 Nov. IET